

DOI: 10.11992/tis.201801025

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.tp.20180420.1029.004.html>

基于 Faster R-CNN 的人体行为检测研究

莫宏伟, 汪海波

(哈尔滨工程大学 自动化学院, 黑龙江 哈尔滨 150001)

摘要: 由于人体行为类内差异大, 类间相似性大, 而且还存在视觉角度与遮挡等问题, 使用人工提取特征的方法特征提取难度大并且难以提取有效特征, 使得人体行为检测率较低。针对这个问题, 本文在物体检测的基础上使用检测效果较好的 Faster R-CNN 算法来进行人体行为检测, 并对 Faster R-CNN 算法与批量规范化算法和在线难例挖掘算法进行结合, 有效利用了深度学习算法实现人体行为检测。对此改进算法进行实验验证, 验证的分类和位置精度达到了 80% 以上, 实验结果表明, 改进的算法具有识别精度高的特点。

关键词: 人体行为检测; 更快速区域卷积神经网络; 在线难例挖掘; 深度学习; 目标检测; 卷积神经网络; 批规范化; 迁移学习

中图分类号: TP181 文献标志码: A 文章编号: 1673-4785(2018)06-0967-07

中文引用格式: 莫宏伟, 汪海波. 基于 Faster R-CNN 的人体行为检测研究[J]. 智能系统学报, 2018, 13(6): 967-973.

英文引用格式: MO Hongwei, WANG Haibo. Research on human behavior detection based on Faster R-CNN[J]. CAAI transactions on intelligent systems, 2018, 13(6): 967-973.

Research on human behavior detection based on Faster R-CNN

MO Hongwei, WANG Haibo

(College of Automation, Harbin Engineering University, Harbin 150001, China)

Abstract: Because of large intra-class difference and large inter-class similarity of human behaviors, as well as problems such as visual angle and occlusion, it is difficult to extract features, especially effective features, using the manual feature extraction method. This results in low accuracy of human behavior detection. To solve this problem, this paper applies a faster region-based convolutional neural network (Faster R-CNN) algorithm, which has a better detection effect, to detect human behavior on the basis of object detection. By combining the Faster-RCNN algorithm with batch normalization algorithm and an online hard example mining algorithm, the deep learning algorithm is effectively utilized to detect human behavior. Experimental results show that the accuracy of classification and position of the improved algorithm exceeds 80%, thereby verifying its high recognition accuracy.

Keywords: human behavior detection; faster R-CNN; OHEM; deep learning; object detection; convolutional neural network; batch normalization; transfer learning

人体行为检测作为人机交互的重要技术始终受到计算机视觉的广泛关注。但在现实环境中存在的一些繁杂背景、人体被物体遮盖、人体动作千姿百态等问题, 使得人体行为检测任务难度较大。

近几年深度学习在图像检测领域获得长足发

展, 各种网络结构层出不穷, 但是所用的训练方式都是随机梯度下降算法, 该算法简单高效, 但是需要训练者自主设置超参数, 调参过程乏味而耗时, 批量规范化 (batch normalization, BN)^[1] 算法是针对该问题提出的, 作为 2016 年以来深度学习领域的重要研究结果, 已成为深度学习算法的重要组成部分。

R-CNN(regions with CNN features)^[2] 模型是应用于物体检测任务的经典算法, 该模型的算

收稿日期: 2018-01-16. 网络出版日期: 2018-04-24.

基金项目: 国家自然科学基金项目 (60035117).

通信作者: 莫宏伟. E-mail: honwei2004@126.com.

法思路是首先读入图像后为输入的图像生成大约 2 000 个类别独立的候选区域,之后使用卷积神经网络从每个候选区域中获取长度相同的特征向量,然后使用支持向量机(SVM)^[3]对每一个候选区域中的物体进行检测分类^[4]。R-CNN 模型采用图像仿射变形计算每一个候选窗口固定大小的卷积神经网络输入,不受窗口形状影响。

在 R-CNN 基础上发展而来的 Fast R-CNN^[5]、Faster R-CNN^[6]等在物体检测方面更为有效。Fast R-CNN 模型基于 R-CNN 模型融合了 SPPNet^[7]网络特点,提高训练和测试速度,同时提高了模型检测精度。Fast R-CNN 模型主要解决了 R-CNN 和 SPPNet 3 方面的问题:测试速度慢、训练速度慢和训练占用空间大。

传统目标检测方法包括有区域建议方法和基于区域的卷积神经网络的方法,但基于区域的卷积神经网络计算量较大,之后通过共享卷积大大降低了深度计算量,Fast R-CNN 可以用非常深的网络实现了接近实时的检测速度,这是在忽略生成区域建议框的基础之上实现的,因此如何降低区域建议框的生成时间成为制约卷积神经网络在目标检测领域应用的一大瓶颈。Faster R-CNN 网络通过 RPN 网络生成区域建议框,结合到 Fast R-CNN 网络进行训练,两者共享卷积层,极大降低了计算复杂度。

1 相关工作

在人体行为检测方面,Ji 等^[8]在研究中将卷积神经网络拓展到三维形式,设计出了 3-D 卷积神经网络(3-D convolutional neural networks,3-D CNN),提出了由 3-D 卷积操作进行视频信息时空特性获取的方法。作者在 KTH 人体行为数据库上采用这一方法进行了测试,获得了 90.2% 的识别正确率。

Karpathy 等^[9]采用 Sports-1M 数据集对卷积神经网络进行训练。在框架级别处理所有视频,使用近似重复查找算法确定了具体视频中包含的重复帧数,实验结果显示了尽管存在着小部分标签错误的干扰,但是网络仍然能够保持着较好的学习状态。

Sun 等^[10]采用慢特征分析方法(slow feature analysis,SFA)来实现人体行为检测任务。慢特征分析方法可以从输入的图像信息中学习不变或微弱改变的特征,研究人员将 SFA 方法与深度学习技术结合起来,用以学习视频数据本身的层次表示,使用具有 3-D 卷积和最大池化操作的双层 SFA

学习结构,将方法扩展到大型输入,并从视频中捕获抽象的结构特征,这一方法在 Hollywood2, KTH 和 UCF Sports 等体行为数据集上的测试也获得了很好的检测效果。

Donahue 等^[11]提出了一个长期递归卷积网络模型。结合深层次的视觉特征提取器(例如 CNN)和模型,可以用于学习识别语言、视觉信息等任务,科研人员将这一方法应用到人体行为检测任务中,并且在 UCF101 数据集上检验了模型的准确性。

Gkioxari 等^[12]提出了利用卷积神经网络进行人体姿势预测和人体行为检测的方法,在 PASCAL VOC 数据集上对这一方法进行了检验,并将其与已有的主要方法进行了对比。

Gkioxari 等^[13]提出了一种根据图像中的情境线索来进行人体行为检测的方法,并取得了较高的识别正确率,此外这一方法还可用于属性分类任务。Khan 等^[14]提出了一种语义金字塔的姿态归一化方法来识别静态图像中的人体行为,并在 PASCAL VOC 2010 和 stanford-40 数据集上进行测试,取得了较高的识别精度。

Christoph 等^[15]提出了一种通用卷积构架,这一架构基于时空特征的乘性交交互对视频中人体行为进行检测。此架构是完全时空卷积,并能够在单一的视频通道进行视频评估。实验表明,该模型在两个标准动作数据集上均具有较好的效果。

Gkioxari 等^[16]通过研究人体部件对动作和属性分类的重要性,提出了一种基于人体部件,利用卷积神经网络进行人体行为检测的方法。实验表明,该方法对人体动作的分类具有较好的效果。

Kar 等^[17]提出了一种新的视频帧合并方法,用于人类行为检测的任务。该方法发现少量的视频帧有可能包含了足够的信息,以此为依据进行视频中的行为分类。并在两个标准数据集上实验,实现了较高的识别效果。

Feichtenhofer 等^[18]将双流卷积神经网络和残差网络 ResNet 进行结合,提出了一种时空域上的人体行为检测方法,该方法依然使用了两个流,运动流和外观流,运动流接收的输入依然是堆叠的多帧光流灰度图片。该方法刷新了 HMDB51 数据库和 UCF101 数据库行为检测的精度。

Herath 等^[19]对过去 10 年的人体行为检测从传统方法到深度学习的方法进行了系统的综述,筹集了新的问题,对人体行为检测的进一步研究指明了方向。

中国科学技术大学邓利群研究团队^[20]提出了

一种以动作检索表为主要依据的识别方法, 该方法提出了将动作信息分割成 5 个片段, 然后针对各片段执行聚类操作和映射操作等, 每种动作的检测任务都需要经过全局检索对比和类型匹配等操作来完成。申晓霞等^[21]提出了基于深度图和 RGB 图像的行为描述算法, 并将其应用到了人体行为检测领域; 王忠民等^[22]将卷积神经网络应用于人体行为检测中, 在应用中将经过简单处理的原始数据直接作为卷积神经网络的输入, 然后由卷积神经网络进行特征获取, 再将得到的特征交由分类器执行分类操作, 该方法对测试人员的行为识别正确率达到了 84.8%, 验证了算法的有效性; 迟元峰等^[23]通过对卷积神经网络和循环神经网络结构的研究分析, 提出了一种新的深度神经网络结构用于检测视频中的人体行为, 实验结果表现出了该方法应用在人体行为检测方面所取得的良好效果; 朱煜等^[24]对近年来人体行为检测方法出现的新进展进行了分析总结和算法综述。

2 改进 Faster R-CNN 算法

BN 层的基本思想很直观: BN 就是通过一定的规范化手段, 把每层神经网络任意神经元输入值的分布强行拉回到均值为 0, 方差为 1 的标准正态分布, 这样可以大大加快训练速度, 同时避免了因参数变化而导致的梯度消失的问题。

如果说批量规范化算法是在网络的结构上对算法进行优化, 那么在线难例挖掘 (online-hard-example-mining, OHEM)^[25]算法则关注如何利用现有的样本对网络进行更有效的训练。

在 Fast R-CNN 算法中提出的 mini-batch 用来进行随机梯度下降时并不具有高效和最优的状态, 而 OHEM 算法可以取得更低的训练损失和更高的平均类别和位置精度 (mAP)。最初的 OHEM

算法与 Fast R-CNN 进行结合使检测算法得到了一定程度的提升, 所以在理论上将 OHEM 算法与 Faster R-CNN 算法进行结合, 可以进一步提高 Faster R-CNN 的检测效果。

简单来说, OHEM 算法就是从 RoI 中选择难例, 而不是简单的采用。难例挖掘一般需要迭代地进行模型的训练和用模型寻找难例, 这些难例无非就是那些分错了并且置信度还很高的样本, 这要是放在网络里面, 这个样本对应的训练损失就应该很大, 因此一个基本思路是利用网络训练损失找到这些难例, 具体方法如下:

1) 对于一张输入图像, 首先利用卷积网络计算出特征图。

2) RoI 网络利用 1) 中计算好的特征图和 RPN 网络得到的 RoI 进行一次前向传播, 最后的 loss 表示 RoI 网络对这个 RoI 检测的好坏程度, 将每一个 RoI 对应的 loss 进行排序, 取前面 B/N 个 RoI 进行反向传播。

3) 因为在同一张图中的 RoI 很多之间相互重叠, 如果一个 RoI 是难例, 那么和它重叠较大的 RoI 很大程度上也是难例, 这样在反向传播的时候就进行了一次重复的梯度计算。为了避免这种情况, 首先根据损失大小进行排序, 然后进行非极大值抑制, 最后再选择 B/N 个 RoI 进行反向传播, 实验中, 非极大值抑制的阈值选取 0.7。

4) 提出的在线难例挖掘算法主要是通过两个 RoI 网络实现的: 只读 RoI 网络 (图 1 上部分所示) 和标准 RoI 网络 (图 1 下部分所示), 只读的 RoI 网络进行前向传播并计算出所有的 RoI 的损失; 之后标准 RoI 样本挖掘模块进行样本挖掘, 得到困难样本并输入到只读 RoI 网络中, 网络只读困难样本进行前向和后向传播, 计算梯度并传播到卷积神经网络。

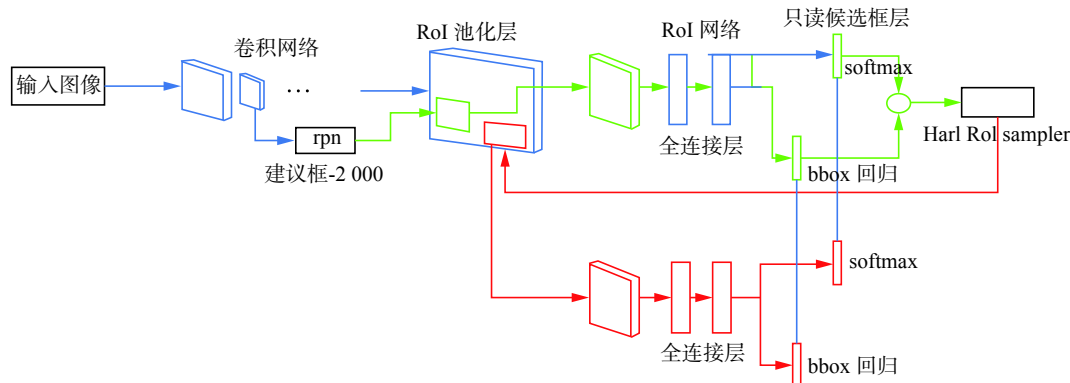


图 1 Faster RCNN_OHEM 算法原理图

Fig. 1 Faster R_CNN_OHEM algorithm schematic

在改进算法中, 首先使用更深的预训练模型 ResNet 进行参数初始化和特征提取。其次使用

BN 算法对输入样本和隐藏层中的每个输入都进行批量规范化处理, 以允许网络使用较大的学习

速率进行训练, 加快网络的训练速度, 具体做法如图2所示。最后使用 OHEM 算法选择难例进行前向和后向传播, 计算梯度并传播到卷积神经网络, 以获得更低的训练损失和更高的类别和位置精度。改进算法的整体流程见图2。

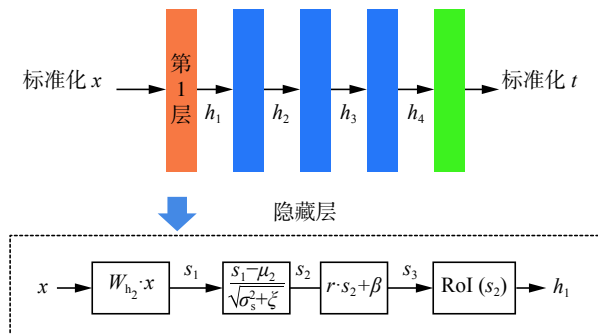


图2 批量规范化处理

Fig. 2 Batch normalization

3 实验过程

3.1 数据集及实验环境

本文实验数据是从 VOC 2012 Action 选取包括 8 个目标类别的人体行为数据集, 3 347 张标记图像, 每类动作大约 400 张。在数据集中, 训练验证集大约占整个数据集的 50%, 测试集大约占整个数据集的 50%, 训练集大约是训练验证集的 50%, 验证集大约占训练验证集的 50%。算法利用 caffe 开源框架, 在 Ubuntu14.04 系统下实现, 所用的计算机配置是, 显卡为 GeForce GTX1080Ti, 内存 11 GB。

3.2 参数选择

使用变学习率策略训练改进的 Faster R-CNN 算法模型, 初始学习率 Base_lr 为: 0.001, 批大小 Batches 为 128, 训练步数 Stepsize 为 6 000, 变学习率参数 gamma 为 0.1。采用了变学习率策略公式:

$$\text{lr} = \text{base}_{\text{lr}} \times \text{gamma}^{\frac{\text{floor} - \text{iter}}{\text{stepsize}}} \quad (1)$$

式中 iter 为当前的迭代次数。

合硬件条件和学习效率之后, 在用 Fast R-CNN 进行人体行为检测时选择 Batch_size 为 128。

3.3 模型训练

Faster R-CNN 模型训练阶段使用训练好的 ResNet 模型参数对 RPN 模型进行初始化, 然后独立地对 RPN 模型进行训练, 之后依旧采用在 ImageNet 上训练好的 ResNet 模型参数对 Fast R-CNN 模型参数进行初始化, 可以看出 RPN 网络与 Fast R-CNN 模型的训练是单独进行的, 两个网络不进行参数共享, 将上一步中 RPN 模型生成的建议窗口作为输入样本对 Fast R-CNN 模型进行训练, 再使用训练后的 Fast R-CNN 模型参数对

RPN 模型再次初始化, 但不更新 RPN 与 Fast R-CNN 共享的卷积层参数, 仅对 RPN 网络特有层参数进行更新, 再进行训练, 此时已实现了两个共享, 之后再对 Fast R-CNN 模型进行训练, 如此交替进行训练。

在训练之前首先修改文件中的相关的参数, 将类型换成人体行为数据集中的标签。将相应输出的类别数改为类别数加一 (类别加上背景)。

4 实验结果与对比分析

测试的结果见表1。表2为 ZF+RPN、VGG+RPN 实验结果和改进算法的实验结果。

表1 改进 Faster R-CNN 算法检测各类别的 AP 值

Table 1 AP of actions of improved Faster R-CN

人体行为	AP
拍照	0.747
打电话	0.786
玩电脑	0.853
阅读	0.735
骑马	0.932
骑车	0.891
跳	0.783
玩乐器	0.696

表2 ZF 和 VGG16 的 mAP

Table 2 mAP for ZF and VGG16

训练	propos als	测试	proposals	mAP/%
RPN+ZF	300	RPN+ZF	300	0.58
RPN+VGG16	300	RPN+VGG16	300	0.68
RPN+OHEM+BN	300	RPN+OHEM+BN	300	0.825

图3为 ZF+RPN、VGG+RPN 和改进算法识别: 8 个类别的平均精度 (AP) 柱状图。

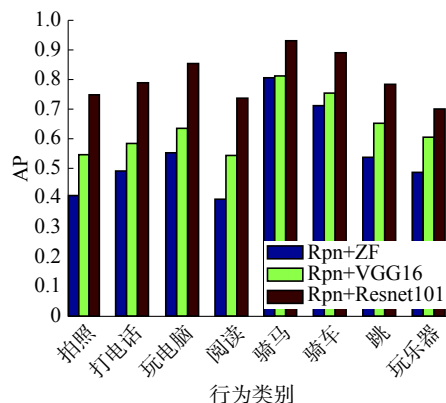


图3 3种方法的 AP 比较

Fig. 3 AP comparison of three methods

在测试数据中随机选取了 12 张人体行为图片进行抽样测试, 测试的结果分别展现在图 4~5 中。图 4 表示未改进算法的测试结果, 从左至右从上至下的人体行为分别为骑马 (1.000)、用电脑 (0.954)、玩乐器 (0.981)、打电话 (0.994)、阅读 (0.997)、骑车 (0.996)、阅读 (1.000)、玩乐器 (1.000)、跳 (0.897)、阅读 (未检测出)、拍照 (未检测出) 和骑车 (未检测出), 其中括号中为检测的位置精确度。图 5 表示改进算法的测试效果, 从左至右从上至下的人体行为分别为骑马 (1.000)、用电脑 (1.000)、玩乐器 (1.000)、打电话 (1.000)、阅读 (0.999)、骑车 (0.996)、阅读 (1.000)、玩乐器 (1.000)、跳 (0.999)、阅读 (0.987)、拍照 (0.934) 和骑车 (0.926), 其中括号中为检测的位置精确度。对比随机选取的测试数据, 图 4 中最后 3 个动作分别为阅读、拍照和骑车, 使用 VGGNet 时算法并未检测出图中的动作。在图 5 中同样的图片、同样的行为, 改进算法对 3 种动作均作出了精确的识别, 并且图片中的动作均比原始算法的识别精确度要高。相比两种算法, 改进的算法不仅能准确地识别出抽样测试图像中所有的人体行为, 并且在精度上有所提升。

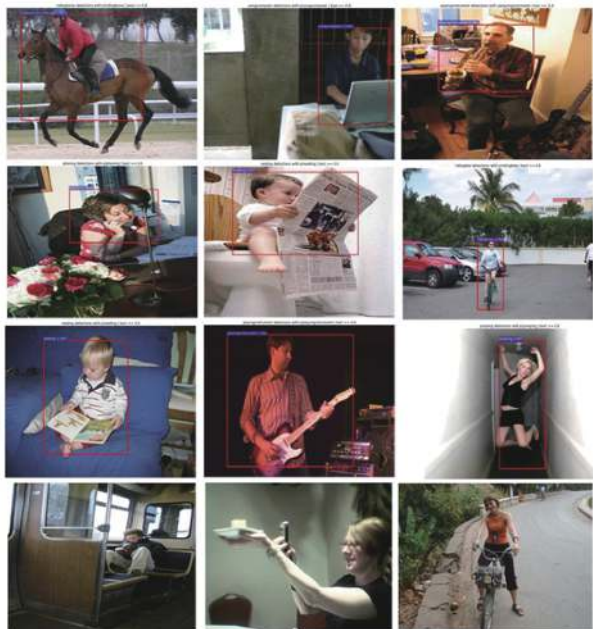


图 4 原始算法检测结果

Fig. 4 The original algorithm detection results.

结合表 2 和图 5 的实验结果以及随机抽样测试的结果可知, 在 8 种不同的人体行为中, 改进的 Faster R-CNN 算法对玩电脑、骑马和骑车这 3 种动作的识别效果较好, 虽然阅读和玩乐器这两种动作的识别效果相比于其他动作效果略差,

但是相对于原始算法识别类别和位置的精度都有着较为明显的改善。并且对于这 8 种动作, 改进 Faster R-CNN 算法识别效果有着显著提升, 平均分类效果和位置精确度均达到 80% 以上, 证明改进算法对人体行为检测任务的有效性。



图 5 改进算法检测结果

Fig. 5 The improved algorithm detection results

5 结束语

Faster R-CNN 算法在物体检测方面有着较好的识别效果, 本文针对人体行为检测关键问题将 OHEM 算法和批量规范化算法与 Faster R-CNN 算法进行结合改进。在改进算法中, 利用 OHEM 算法的优势可以有效地识别出静态图像中存在的小样本难例, 对于人体行为识别整体的识别正确率都有所改善。今后将在通过搜集更多测试样本, 增加类别, 来测试所提出方法的有效性。

参考文献:

- [1] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//Proceedings of the 32nd International Conference on International Conference on Machine Learning. Lille, France, 2015: 448–456.
- [2] GIRSHICK R, DONAHUE J, DARRELL T, et al. Region-based convolutional networks for accurate object detection and segmentation[J]. IEEE transactions on pattern analysis and machine intelligence, 2016, 38(1): 142–158.
- [3] 李航. 统计学习方法[M]. 北京: 清华大学出版社, 2012:

- 36–58.
- LI Hang. Statistical learning method[M]. Beijing: Tsinghua University Press, 2012: 36–58.
- [4] 张文达, 许悦雷, 倪嘉成, 等. 基于多尺度分块卷积神经网络的图像目标识别算法[J]. 计算机应用, 2016, 36(4): 1033–1038.
- ZHANG Wenda, XU Yuelei, NI Jiacheng, et al. Image target recognition method based on multi-scale block convolutional neural network[J]. Journal of computer applications, 2016, 36(4): 1033–1038.
- [5] GIRSHICK R. Fast R-CNN[C]//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 1440–1448.
- [6] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada, 2015: 91–99.
- [7] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[C]//Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland, 2014: 346–361.
- [8] JI Shuiwang, XU Wei, YANG Ming, et al. 3D convolutional neural networks for automatic human action recognition[P]. USA: 8345984, 2013.
- [9] KARPATY A, TODERICI G, SHETTY S, et al. Large-scale video classification with convolutional neural networks[C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA, 2014: 1725–1732.
- [10] SUN Lin, JIA Kui, CHAN T H, et al. DL-SFA: deeply-learned slow feature analysis for action recognition[C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA, 2014: 2625–2632.
- [11] DONAHUE J, HENDRICKS L A, GUADARRAMA S, et al. Long-term recurrent convolutional networks for visual recognition and description[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA, 2015: 2625–2634.
- [12] GKIOXARI G, HARIHARAN B, GIRSHICK R, et al. R-CNNs for pose estimation and action detection[J]. Computer science, 2014, 12(8): 1221–1229.
- [13] GKIOXARI G, GIRSHICK R, MALIK J. Contextual action recognition with R*CNN[C]//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 1080–1088.
- [14] KHAN F S, VAN DE WEIJER J, ANWER R M, et al. Semantic pyramids for gender and action recognition[J]. IEEE transactions on image processing, 2014, 23(8): 3633–3645.
- [15] FEICHTENHOFER C, PINZ A, WILDES R P. Spatiotemporal multiplier networks for video action recognition[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017: 7445–7454.
- [16] GKIOXARI G, GIRSHICK R, MALIK J. Actions and attributes from wholes and parts[C]//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 2470–2478.
- [17] KAR A, RAI N, SIKKA K, et al. Adascan: adaptive scan pooling in deep convolutional neural networks for human action recognition in videos[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017: 5699–5708.
- [18] FEICHTENHOFER C, PINZ A, WILDES R P. Spatiotemporal residual networks for video action recognition [C]//Proceedings of the 30th Conference on Neural Information Processing Systems. Barcelona, Spain, 2016: 3468–3476.
- [19] HERATH S, HARANDI M, PORIKLI F. Going deeper into action recognition: a survey[J]. Image and vision computing, 2017, 60: 4–21.
- [20] 邓利群. 三维人体动作识别及其在交互舞蹈系统上的应用[D]. 合肥: 中国科学技术大学, 2012.
- DENG Liqun. 3D mocap data recognition and its application on interactive dancing game[D]. Hefei: University of Science and Technology of China, 2012.
- [21] 申晓霞, 张桦, 高赞, 等. 基于深度信息和 RGB 图像的行为识别算法[J]. 模式识别与人工智能, 2013, 26(8): 722–728.
- SHEN Xiaoxia, ZHANG Hua, GAO Zan, et al. Behavior recognition algorithm based on depth information and RGB image[J]. Pattern recognition and artificial intelligence, 2013, 26(8): 722–728.
- [22] 王忠民, 曹洪江, 范琳. 一种基于卷积神经网络深度学习的人体行为识别方法[J]. 计算机科学, 2016, 43(11A): 56–58, 87.
- WANG Zhongmin, CAO Hongjiang, FAN Lin. Method on human activity recognition based on convolutional

neural networks[J]. Computer science, 2016, 43(11A): 56–58, 87.

- [23] 迟元峰, 顾敏. 基于深度学习的人体行为识别研究[J]. 工业控制计算机, 2017, 31(1): 104–105.

CHI Yuanfeng, GU Min. Human action recognition based on deep learning[J]. Industrial control computer, 2017, 31(1): 104–105.

- [24] 朱煜, 赵江坤, 王逸宁, 等. 基于深度学习的人体行为识别算法综述[J]. 自动化学报, 2016, 42(6): 848–857.

ZHU Yu, ZHAO Jiangkun, WANG Yi'ning, et al. A review of human action recognition based on deep learning[J]. Acta automatica sinica, 2016, 42(6): 848–857.

- [25] SHRIVASTAVA A, GUPTA A, GIRSHICK R. Training region-based object detectors with online hard example mining[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016: 761–769.

作者简介:



莫宏伟, 主要研究方向为人工智能、类脑计算、智能机器人。承担完成国家自然科学基金、国防预研等项目 17 项。中国人工智能学会自然计算与数字城市专业委员会副主任, 黑龙江省生物医学工程学会理事。中国生物医学工程学会高级会员。中国计算机学会高级会员。International Journal of Swarm Intelligence Research、《电子学报》编委。IEEE Tran on Industrial Informatics 2018 专刊《医疗卫生中的大数据处理》副主编。发表学术论文 70 余篇。出版专著 6 部, 授权发明专利 7 项。



汪海波, 男, 1990 年生, 硕士研究生, 主要研究方向为深度学习。

2019 年第 11 届机器学习和计算国际会议 2019 11th International Conference on Machine Learning and Computing (ICMLC 2019)

会议网站: <http://www.icmlc.org>

2019 年 2 月 22-24 日 珠海

文章出版

被录用的文章将出版到 ACM 会议论文集, 并提交 EI 核心检索, Scopus 检索。

ICMLC 2017 ACM 会议论文集 (ISBN: 978-1-4503-6353-2) 已经提交给 ACM, 很快将会被 ACM Digital Library 收录, 被 Ei Compendex 和 Scopus 检索。ICMLC 2017 ACM 会议论文集 (ISBN: 978-1-4503-4817-1) 已被 ACM Digital Library 收录, 已被 Ei Compendex 和 Scopus 检索。

会议主讲人:

Prof. Tianrui Li, Southwest Jiaotong University, China

Prof. David Greenhalgh, University of Strathclyde, UK

Assoc. Prof. Lin Huang, Metropolitan State University of Denver, USA

投稿方式

1. 全文 (报告和出版);

2. 摘要 (仅报告)。

请登录网上投稿系统 (<https://www.easychair.org/conferences/?conf=icmlc2019>) 或直接发邮件到 icmlc@asr.org 进行投稿。

联系方式

会议秘书: 苗老师;

邮箱: icmlc@asr.org;

电话: +86-13258-11111-7.